application is placed on line. The amendments to paragraphs 81 and 88 delete references to a Figure not included in the application as filed.

The amendments to claims 1 and 21 add a recitation that the mutations alter the sequence of the encoded protein from the sequence encoded by the wild-type gene. This is supported throughout the specification, including Figures 4B and C, and page 8, paragraph 22. The amendments to claims 12 and 29 clarify the wording, as suggested by the Examiner. The amendment to claim 25 adds a recitation that the primers specifically amplify the EtaA gene or a portion thereof. Primers for specifically amplifying the EtaA gene or portions thereof are supported throughout the specification, including paragraphs 62 and 77.

Claims 34-48 track original claims 1-5, 8-12, 21, 22, 25, 28 and 29 as amended herein, but with the recitation of claim 16 incorporated into the independent claims.

III. The Office Action

The Action rejects the pending claims on several grounds. Applicants amend in part and traverse. For the Examiner's convenience, the rejections are discussed below in the order in which they are presented in the Action.

A. Restriction Requirement

Applicants note with appreciation the Examiner's acknowledgement that claim 1 is a linking claim and that claims 2-4 will be examined if the linking claim is allowed. Applicants respectfully disagree with the Action's other contentions regarding the restriction and note that the rules permit a further challenge to be raised to the restriction at any time during the prosecution.

B. Objection to the Specification

Citing MPEP § 608.01, the Action objects to the specification as containing an embedded hyperlink at page 10.

A review of MPEP § 608.01 indicates that the Office's concern is that, when an application is placed on the PTO's website, embedded hyperlinks will become active and would take readers to websites not under the PTO's control. The citation at page 10 of the specification has been edited so that it provides the same information but will not be active as a hyperlink when placed on-line. A second hyperlink was noted in the preparation of this Amendment and has also been edited so that it will not create an active hyperlink when placed on the PTO's website.

C. Rejection of the claims under §112, first paragraph

Claims 1-5, 8-12, 16, 21, 22, 25, 28, and 29 are rejected under 35 U.S.C. § 112, first paragraph, as not enabled. The Action concedes that the specification is enabled for methods of determining the ability of a *Mycobacterium tuberculosis* ("Mtb") bacterium to oxidize a carbonyl in the drugs ethionamide ("ETA"), thiacetazone ("TA") and thiocarlide ("TC") comprising detection one of the mutations exemplified in the specification in the Mtb EtaA gene. The Action contends, however, that the claims are "broadly drawn" to any mutation in an Mtb EtaA gene, and the ability of an Mtb to oxidize any thioamide or thiocarbonyl. The Action states that the specification does not enable methods of

> determining the ability of an *Mycobacterium tuberculosis* bacterium to oxidize any thioamide or any thiocarbonyl, said method comprising detecting any mutation in the EtaA gene in said bacterium, wherein detection of the mutation is indicative of decreased ability to oxidize any thioamide or any thiocarbonyl.

Action, at page 4. Applicants amend in part and traverse. For clarity, the discussion below attempts to parse out and respond individually to the various arguments set forth in the Action.

1. The rejection is based on a misunderstanding of the invention and on a misreading of the specification's data

The Action states that Figure 4C of the specification teaches "that several strains with mutations [in the EtaA gene] were conferred with drug susceptibility instead of resistance." Action, at page 6 (emphasis added). The Action therefore fundamentally misunderstands the invention and misreads the data presented.

As an initial matter, Applicants note that it appears that the Action may mistake the front-line drug isoniazide ("INH") with the thioamide or thiocarbonyl drugs recited by the claims. INH contains an oxygen double bonded to a carbon, not a sulfur double bonded to a carbon. It therefore contains a **carbonyl** group, not a **thiocarbonyl,** as recited by the claims under examination. (For the reader's convenience, Applicants note that the composition of INH and of ethionamide are set forth in the DeBarber reference (DeBarber et al., Proc Natl Acad Sci USA 97(17):9677-9682 (2000) cited by the Action against the application, at page 9681, Figure 5).

The statement that strains with mutations "were conferred with drug susceptibility" shows a fundamental misunderstanding of the invention: it is a discovery of the present invention that organisms with mutations in the EtaA gene are resistant to thioamide and thiocarbonyl drugs, while those with the unmutated, wild-type gene are susceptible to them. For example, the specification teaches, at page 7, lines 28-30, "that it has now been discovered that this gene [EtaA] confers upon *Mycobacteria* the ability to activate thioamide and thiocarbonyl drugs from their prodrug form to their active drug form." (Emphasis added.) Thus, the product of the wild-type EtaA gene activates ETA from a prodrug (inactive) form to the active therapeutic drug. In contrast, organisms with mutations in the gene are resistant to ETA, and to related drugs. In this regard, the specification teaches: "organisms with mutations in the EtaA gene are resistant not only to ETA, but also to two other thioamide compounds also used as second-line drugs." Specification, at page 8, lines 23-25. Thus, contrary to the Action's statement, organisms

with the wild-type EtaA gene are susceptible to ETA; mutations in the gene confer drug resistance, not drug susceptibility.

Second, the Action misreads the data presented in Figure 4C as teaching that mutations in the gene confer drug susceptibility to thioamide and thiocarbonyl drugs. A review of Figure 4C shows that all 11 patient isolates showing at least some resistance to ETA (isolates $AS1TA_R$ to $AS11TA_R$) contained a mutation in the EtaA gene. All 11 of these isolates also showed at least some resistance to thiacetazone and 10 of the 11 showed resistance to thiocarlide. Thus, the presence of a mutation in the EtaA was shown to be correlated with a high degree of predictability to resistance to thioamide and thiocarbonyl drugs. As noted above, the front-line agent isoniazide (INH) contains a carbonyl, not a thiocarbonyl.

The Action appears to mistake the three patient isolates that were susceptible to ETA (isolates $AS12TA_R$ to $AS14TA_R$) as showing that mutations in the EtaA gene conferred susceptibility to ETA. Reading across the rows of Figure 4C for these isolates, however, reveals that none of them had a mutation in the EtaA gene.

Accordingly, the Action's argument is based on a misreading of the data presented in the Figure. The rejection should therefore be reconsidered. Applicants submit that, upon reconsideration, it should be withdrawn.

2. The rejection is based in part on a series of 5 incorrect statements to the effect that the specification does not teach that mutations in the EtaA gene are correlated with resistance to thioamide and thiocarbonyl drugs

At page 6, the Action presents a list of teachings presumably absent from the specification. Specifically, the Action asserts at page 6 that the specification does not teach: (i) "that every mutation in the EtaA gene confers a drug resistance to the bacterium as a result of its inability to oxidize thiocarbonyl groups," (ii) "the way in which each drug, ETA, TA, TC, or INH [sic] differs," (iii) "the characteristic shared between all

drugs responsible for conferring the phenotype", (iv) "the other mutations in the EtaA gene that would be indicative of either the drug resistant phenotype, or the drug sensitive phenotype," and (v) "a common property represented within each mutation that is responsible for the resulting phenotype of drug resistance." Applicants amend in part and traverse. For clarity, the 5 contentions are addressed individually or in smaller, related groups below.

(a) The specification teaches that mutations in the EtaA gene will result in drug resistance (Response to contentions (i) and (iv))

Applicants turn first to contention (i), that the specification does not teach that every mutation in the EtaA gene confers a drug resistance to the Mtb bacterium, and contention (iv) that the specification has not indicated other alleged mutations that would be indicative of either a drug resistant phenotype of a drug sensitive phenotype.

The specification teaches that it was known in the art that Mtb "has an extremely low rate of synonymous mutations, that is, that the organism has few, if any, random mutations which do not have a functional effect. E.g., Sreevatsan, S., et al., Proc Natl Acad Sci USA 94(18):9869-74 (1997)." Specification, at page 8, lines 10-14. Applicants note that the Sreevatsan reference was not only available in the literature before the priority date, and was therefore known to those of skill in the art, but also is incorporated by reference into the specification (see, page 30, lines 28-30) and is therefore legally part of its teachings. A copy of Sreevatsan is attached for the Examiner's convenience.

Sreevatsan et al. studied the nucleotide sequence diversity of 26 genes from 715 Mtb isolates from around the world, as well as over 100 isolates of the three closely related Mycobacteria known as the "*M. tuberculosis* complex." Sreevatsan, at page 9869, right hand column and Table 1. The authors found that "[c]ompilation of the two megabases of sequence data for the 26 genes revealed that greater than 95% of

nucleotide substitutions caused amino acid replacements or other mutations in gene regions linked to antibiotic resistance," (page 9870, bottom right), and that "greater than 95% of nucleotide changes were directly associated with antibiotic resistance." (*Id.*, page 9873, left column, second full paragraph.) The authors conclude that "[t]he lack of allelic diversity means that when amino acid polymorphisms, or regulatory region nucleotide variation are observed, there should be strong suspicion that the variation has functional consequences, such as antibiotic resistance." *Id.*, at page 9872, bottom right hand column.

Accordingly, the art as of the priority date taught that there was a high probability that any nucleic acid substitutions or other mutations in Mtb genes such as EtaA would result in antibiotic resistance. The Action presents no data or evidence to the contrary. Thus, the person of skill in the art would expect that, as taught in the specification at page 8, lines 14-15, any mutation in the EtaA gene would result in increased antibiotic resistance. This expectation would be confirmed by the specification, which presents data showing that every mutation found in the EtaA gene in patient isolates (a total of nine different mutations) resulted in increased resistance. See, specification at page 8, lines 15-20 and Figure 4C. No silent mutations were found.

Accordingly, Applicants respectfully maintain that the specification and the art support the conclusion that every mutation in the EtaA gene will reduce the ability of a Mtb organism to oxidize a thioamide or thiocarbonyl drug, and therefore increase the resistance of the organism.

While this answers contention (i), it also answers contention (iv). Every mutation in the gene is expected to result in a drug resistant phenotype. Contention (iv) appears to be based on the Action's misreading of Figure 4C as showing some EtaA mutants as being susceptible to ETA. This misreading of the Figure is discussed, and refuted, in Section III C 1, above.

To expedite prosecution, however, claims 1 and 21 have been amended to recite that the mutations result in an alteration of the amino acid sequence of the protein

encoded by the mutated gene compared to SEQ ID NO:2, the amino acid sequence encoded by the wild-type gene. The amended claims eliminate the possibility of encompassing mutations of a nucleotide in codon which, because of the degeneracy of the genetic code, would not result in a change in the amino acid encoded by the codon. Applicants respectfully observe that the specification teaches no fewer than nine such mutations (two frameshift mutations and seven amino acid substitutions) that result in a decreased ability of the bacterium to oxidize a thioamide or thiocarbonyl.

(ii) The specification shows a common property represented by the mutations (Response to contentions (ii), (iii), and (v)).

The Action further contends that the specification does not teach the way ETA, TA, TC and INH differ (contention (ii)), show a characteristic shared by ETA, TA, and TC (contention (iii)), or show a common property represented within each mutation that is responsible for the phenotype of drug resistance (contention (v)).

As an initial matter, the Action incorrectly includes INH as one of the drugs Applicants are to explain. As noted in Section C 1, above, INH contains a **carbonyl** group, not a **thiocarbonyl**, as recited by the claims under examination. Ability of an Mtb bacterium to metabolize INH is therefore not at issue in this proceeding.

Applicants note that the specification shows the metabolism of ETA (see, specification, at page 25, Example 5, to page 27, end of Example 6). The specification further shows that two other thioamide or thiocarbonyl-containing drugs in addition to ETA were selected to elucidate the clinical relevance of EtaA-mediated resistance to thiocarbonyl-containing drugs as a class (specification, at page 28, lines 12-15). The specification teaches that 11 of 11 patient isolates of persons resistant to thiacetazone and having mutations in the EtaA gene were cross resistant to ETA. *Id.*, at lines 15-18, and 23-25, and Figures 4B and C. Moreover, the specification discloses that 10 out of 11 of

these individuals were resistant to thiocarlide even though they had never been treated with it. *Id.*

Persons of skill in the art were both advised of the metabolic pathway on which the EtaA gene product acts and of evidence supporting that mutations in the EtaA gene interfere with a feature common to the action of thioamide and thiocarbonyl drugs. Given the high degree of correlation shown in the specification between mutations in the EtaA gene and resistance to three exemplar thioamide and thiocarbonyl drugs, Applicants respectfully note that there is predictability that the method of the invention will apply with respect to other thioamide and thiocarbonyl drugs. Applicants respectfully note that this is all that is necessary to enable others to make and use the invention as claimed. They are aware of no provision of the patent statute that requires that they also show exactly how the mutations interfere with drug action to achieve the effect.

Similarly, with respect to contention (iii) and (ii), Applicants respectfully repeat that persons of skill in the art were both adequately advised of the metabolic pathway on which the EtaA gene product acts and of evidence supporting that mutations in the EtaA gene interfere with a feature common to thioamide and thiocarbonyl drugs. This is sufficient to provide predictability that the method of the invention will apply with respect to other thioamide and thiocarbonyl drugs. Applicants respectfully note that this is all that is necessary to enable others to make and use the invention as claimed. They are aware of no provision of the patent statute that requires that they also show how, in the words of the contention, each of the drugs differs (or, perhaps more precisely, how their mechanism of action differs).

c. The Action's contentions that the drugs ETA, TA, and TC "confer different phenotypes to the bacterium" and "confer varying degrees of drug resistance" reverse the teachings of the specification

The Action's contention that the invention is claimed more broadly than is supported rests in part on the assertion that: "the specification teaches that each of the drugs, ETA, TA, and TC confer different phenotypes to the bacterium," Action, at page 6, and does not teach "how or why such supposedly similar thiocarbonyl-containing antituberculosis medications confer different degrees, if at all, of drug resistance." *Id:*

Applicants respectfully note that the specification does not teach that the various anti-tuberculosis drugs discussed "confer" a phenotype on any organism: the phenotype of the organism is whether it is susceptible or resistant to a particular drug. What the specification does teach is that the organism's genotype, with respect to the EtaA gene, predicts the organism's phenotype: a mutation in the EtaA gene is correlated with resistance to thioamide and thiocarbonyl drugs. The Action therefore appears to misunderstand, or to reverse, the specification's teachings. Applicants surmise that the assertion is grounded in the Action's misreading of Figure 4C as showing that one of the patient isolates having a mutated EtaA genes was susceptible to ETA. As shown in a preceding section, every patient isolate in which an EtaA gene mutation was found was resistant to ETA. The assertion therefore does not support the Action's contention that the specification does not enable the claims as presented.

Applicants also surmise that the Action may be trying to refer to the fact that Figure 4C shows that three patient isolates (isolates $AS12TA_R$ to $AS14TA_R$) that did not have a mutation in the EtaA gene were susceptible to ETA gene, but resistant to TA or TC, or both. Applicants respectfully note that there is nothing in the claims or in the specification that asserts that mutations in the EtaA are the only way organisms can achieve resistance to drugs or which would prevent other mechanisms of resistance from being present.

Similarly, the Action's assertion that the specification does not teach "how or why such supposedly similar thiocarbonyl-containing antituberculosis medications confer different degrees, if at all, of drug resistance" misunderstands or reverses the invention. The medications, of course, do not confer drug resistance. Applicants surmise that the Action was trying to state that the specification does not make clear how mutations in the gene conferred the different degrees of resistance to ETA, TA and TC shown in Figure 4C. Applicants respectfully note that the claims under examination require only a decrease in ability to oxidize a thioamide or thiocarbonyl; they do not recite that the decrease in ability be the same with respect to each drug. Even assuming this is what the Action was trying to articulate, therefore, the Action would be rejecting the claims over a recitation that is neither present nor required. Applicants respectfully point out that the specification needs only to support the recitations of the claims. Nothing in the patent statutes require that they also enable recitations not present in the claims or of only scientific interest.

As pointed out above, Sreevatsan et al. taught, before the priority date, that "greater than 95% of nucleotide changes were directly associated with antibiotic resistance." (see, Sreevatsan, at page 9873, left column, second full paragraph.), and the claims have been amended to recite only mutations that alter the amino acid sequence of the encoded protein from SEQ ID NO:2, the amino acid sequence encoded by the wild-type gene. Accordingly, persons of skill in the art would expect that such mutations in the EtaA gene would be correlated with drug resistance.

d. The Action's contention that one cannot tell which mutations will cause resistance compared to those that will cause susceptibility is based on a misreading of the data.

At page 7, the Action refers to Figure 4C and argues that "one cannot readily anticipate which of the mutations within the gene . . . actually result in the inability to oxidize thiocarbonyl groups and that would be associated with a patient that is

resistant . . . as opposed to those frameshifts or polymorphisms that result in drug sensitivity." Applicants traverse.

This argument is based on a misreading of Figure 4C. Every frameshift and single nucleotide polymorphism reported in Figure 4C was correlated to at least some resistance to ETA. In 11 out of 11 cases, the mutations were also correlated in resistance to TA, and in 10 out of 11 cases, were correlated with resistance to TC as well. The specification thus shows an extraordinarily high degree of predictability between mutations in the EtaA gene and resistance to thioamide and thiocarbonyl drugs. Applicants respectfully point out that the 10 isolates resistant to TC were from patients that had never been treated with TC (see, specification at page 28, lines 15-19). Thus, detection of mutations in the EtaA gene showed very high correlation with the ability to predict cross resistance to a drug with which the organism had not yet been challenged. In contrast, three patient isolates found to be susceptible to ETA "were mutation free in the EtaA/EtaR and intergenic regions." Specification, at page 8, lines 17-20, and Figure 4C. *See also*, Figure 4C at rows for isolates AS12TA_R to AS14TA_R.

Thus, the methods of the invention are correlated with a high degree of predictability in determining whether or not a Mtb organism is resistant to thioamide and thiocarbonyl drugs. (Applicants note again that the Action appears to incorrectly classify isoniazid as a thiocarbonyl drug. It is a carbonyl drug, and not within the recitation of the claims.) The rejection is based on a misreading of the data presented in the specification. The rejection should be reconsidered and, upon reconsideration, withdrawn.

e. The "post filing date art" cited by the Action does not show unpredictability of the invention

The Action cites DeBarber et al., Proc Natl Acad Sci USA 97(17):9677-9682 (2000) ("DeBarber") as "post filing date art." The Action quotes DeBarber as teaching that the "extensive cross-resistance among these compounds predicts multiple

overlapping mechanisms of resistance" and that "such considerations complicate appropriate drug treatment." Action, at page 7.

Applicants note that the DeBarber reference is of course the inventors' own work and its text and information are the basis of the present application. The precise passage cited by the Action as "post filing date art" is in fact set forth in the specification at page 30, paragraph 92. It is therefore not "post filing date art," but part of the disclosure of the invention itself.

The Action cites the portion of the passage quoted above, and then states: "Thereby, the scope of the claims do not bear a reasonable correlation to the scope of enablement." Applicants respectfully note that the Action goes directly to this conclusion without providing any analysis connecting the passage cited to the alleged lack of reasonable correlation. In fact, the Action alters the sense of the passage by taking it out of context. The complete passage reads:

> The extensive cross-resistance among these compounds predicts multiple overlapping mechanisms of resistance among clinically used antituberculars: target associated between INH and ETA, and activation-associated between ETA, thiacetazone, and thiocarlide. Such considerations complicate appropriate drug therapy for the treatment of multidrug-resistant tuberculosis and <u>these results provide an important tool to help understand and quickly characterize the resistance mechanisms operating in a single patient, which may prove vital to a positive outcome</u>.

Specification, at page 30, lines 17-23 (emphasis added). The underlined portion of the passage indicates the inventors contemplated that the discovery of the association between mutations in the EtaA gene and cross resistance to thioamide and thiocarbonyl drugs would "provide an important tool to help understand and quickly characterize the resistance mechanisms operating in a single patient."

Read in context, therefore, the passage supports, rather than undercuts, the predictability of determining cross resistance to thioamide and thiocarbonyl-containing

drugs by detecting mutations in the EtaA gene.[1] The rejection is therefore founded on an incorrect premise. Applicants respectfully request that the rejection be reconsidered and, upon reconsideration, withdrawn.

f. Application of the *Wands* criteria to the claims under examination show that undue experimentation is not required to practice the invention

The Action further argues that the specification has not taught "an association between these mutations [e.g., the ones set forth in Figure 4B] and the actual effect they have on the bacterium's ability to oxidize [a thioamide or thiocarbonyl drug] and therefore on its potential for resistance. Such random trial by error experimentation is considered to be undue." Action, at pages 7-8, bridging sentences.

The rejection is therefore predicated on the position that, because some experimentation might be required to determine the effect of any particular mutation, it would be "undue." This is not correct. In fact, an invention can require considerable experimentation to practice so long as that experimentation is routine in the art. Whether or not experimentation is undue is determined under criteria that were articulated by the Federal Circuit in *In re Wands,* 8 USPQ2d 1400, 1404 (Fed. Cir. 1988). The Action correctly cites some of the *Wands* criteria, but leaves others out. *Wands* articulates eight criteria.

The full list of criteria articulated in *Wands* is as follows: (1) the breadth of the claims; (2) the nature of the invention; (3) the state of the prior art; (4) the level of one of ordinary skill; (5) the level of predictability in the art; (6) the amount of direction provided by the inventor; (7) the existence of working examples; and (8) the quantity of experimentation needed to make or use the invention based on the content of the

---

1 Applicants note that the copy of DeBarber that was provided with the Action has circles around "S"s in the column under isoniazid (INH). Applicants again point out that INH contains a carbonyl, not a thiocarbonyl (*see*, structure of INH shown in DeBarber, at page 9681, Figure 5), and is therefore not encompassed by the claims under examination. The circles around the "S"s in the column for INH further evidence that the rejection is founded on an incorrect characterization of INH.

disclosure. Application of the *Wands* criteria to the present claims shows that persons of skill in the art were fully enabled to use the invention as claimed without undue experimentation:

Factor 1: The Breadth of the Claims.

The Action contends that claims 1-5, 8-12, 16, 21, 25, 28, and 29 "are broadly drawn to method of determining the ability of a *M. tuberculosis* bacterium to oxidize any thioamide or any thiocarbonyl" by detecting "any mutation in the EtaA gene." Action at page 5.

Applicants respectfully maintain that the scope of the claims is appropriate to the teachings of the specification. The claims are not broadly drawn to determining the resistance of an Mtb to any drug of any class, from detecting a mutation in any gene. They are drawn to a specific gene, the function of which is disclosed for the first time in the specification. They are further drawn to two specific classes of drugs. The studies set forth in the specification show a high correlation between mutations in the gene and resistance to these classes of drugs. And, while the claims do recite that the resistance can be do to any mutation in the EtaA gene (or, as now claimed, any mutation that results in an alteration of the amino acid sequence of the encoded protein, compared to SEQ ID NO:2), this is supported by (a) the fact that every one of nine separate mutations in the gene were found to result in increased resistance to drugs of the claimed classes and (b) the knowledge in the art that greater than 95% of mutations in Mtb genes are known to result in antibiotic resistance.

Accordingly, Applicants maintain that the scope of the claims is appropriate in light of the teachings of the specification.

Factor 2: The nature of the invention.

The invention is related to the discovery of the function of the EtaA gene and of its relation to resistance to thioamide and thiocarbonyl drugs.

The determination of the gene's function and results reported in the specification regarding the ability to predict sensitivity or resistance to these drugs were considered important enough to warrant publication in the Proceedings of the National Academy of Sciences, in the DeBarber paper cited by the Action *against* the application. Applicants note that PNAS is one of the most prestigious and widely followed scientific journals. Applicants submit that the scope of the claims is warranted given the importance of the invention to determining appropriate therapy for persons infected with Mtb resistant to the front-line drug, isoniazid.

Factor 3: The state of the prior art.

Thioamide and thiocarbonyl drugs for treating tuberculosis are, of course, already known in the art. Moreover, the Sreevatsan article shows that the art taught that greater than 95% of mutations in genes in Mtb are associated with functional effects, including antibiotic resistance. Combining this prior art teaching with the findings in the specification that the EtaA gene is associated with the oxidation of thioamide and thiocarbonyl drugs, one of skill in the art would predict that there was a greater than 95% chance that any mutation in the EtaA gene would result in increased antibiotic resistance.

Factor 4: The level of one of ordinary skill

Persons of skill in this art typically possess M.D.s or Ph.D.s. Therefore, the level of skill in the art is very high indeed, and the amount of guidance that the practitioner needs to be provided is correspondingly reduced. As the Office has recognized, there is an inverse correlation between the level of skill of those in the art and the amount of guidance that must be provided to them.

Factor 5: The level of predictability in the art

Substantial predictability can be found in the relevant field. As indicated in the specification and discussed in the preceding sections, 11 out of 11

patient isolates resistant to TA that had a mutation in the EtaA gene were also resistant to ETA. Moreover, 10 of the isolates were also resistant to TC, even though the patients had never been treated with that drug. Therefore, the specification shows a very high correlation between the presence of mutations in the EtaA gene and resistance to thioamide and thiocarbonyl drugs. (As also discussed in more detail above, some of the Action's basis for discounting this correlation is based on the incorrect belief that isoniazid (INH) is a thiocarbonyl drug, which it is not.) The art shows that greater than 95% of mutations in genes in Mtb are associated with functional effects.

Given that the EtaA gene is shown in the specification to be involved in susceptibility to thioamide and thiocarbonyl drugs, persons in the art would expect that there would therefore be a high probability that mutations in the gene would be associated with increased resistance to antibiotics of those types. Finally, the specification shows that both frameshift mutations and single amino acid substitutions, nine separate mutations in all, all resulted in increased resistance to thioamide and thiocarbonyl drugs. The Action provides no example of a single mutation that does not. Accordingly, there is considerable predictability that the invention would work as claimed.

Factor 6: The amount of direction provided by the inventors

The specification provides ample guidance as to how to make and use the invention as claimed. For example, the specification provides considerable information on how to amplify the EtaA gene (page 18, line 23 to page 20, line 13), and primers that can be used in such amplifications (page 20, lines 4-13), the labeling of probes (page 20, line 15 to page 21, line 17). The application further teaches how to detect mutations in the EtaA gene by, for example, cleavase fragment length polymorphism (page 13, lines 13-19), single stranded conformation polymorphism (page 13, lines 20-32), temperature

modulation heteroduplex chromatography (page 13, line 33 to page 14, line 6), hybridization of probes (page 13, lines 7-14), oligonucleotide ligation assay (page 16, lines 3-10), and mass spectroscopy (page 16, lines 17-22). Further, the specification teaches hw to detect functional changes caused by the mutations by screening for inhibition of monooxygenases (page 16, line 23, to page 17, line 2) and by detecting truncated forms of the protein encoded by the gene (page 17, lines 3-17). Further, the specification teaches detecting mutations by culturing Mt and determining whether or not a metabolic product of EtaA gene action is present (page 17, line 18, to page 18, line 2), as well as by the detection of the metabolic product (page 18, lines 3-20).

The specification therefore provides considerable guidance to the practitioner. Moreover, as noted above, there is an inverse correlation between the level of skill of the practitioner and the amount of guidance needed to enable the invention. Given the high degree of skill of practitioners in this art, the specification provided all the guidance necessary to practice the invention as claimed.

Factor 7: The existence of working examples

The specification provides no fewer than nine examples of mutations in the EtaA gene. The specification shows that 11 patient isolates each having one of these mutations showed decreased ability to oxidize the thioamide ETA, 11 of 11 showed decreased ability to oxidize the thioamide TA, and 10 of 11 showed decreased ability to oxidize the thiocarbonyl TC. Moreover, none of the patients had ever been treated with TC, showing the predictive ability of the claimed methods. Although the Action provides no example of any mutation in the EtaA gene that is not associated with a reduced ability to oxidize a thioamide or thiocarbonyl, it contends that the specification is only enabling for the specific mutations tested. Given the teachings of Sreevatsan et al. that "greater than 95%

of nucleotide changes [in two megabases of sequence data on 26 genes of Mtb and organisms of the *M. tuberculosis* complex] were directly associated with antibiotic resistance" Sreevatsan, supra, at page 9873, left hand column, Applicants submit that the results obtained with respect to the nine mutations reported in the specification demonstrate that it is predictable that mutations in the EtaA gene reduce the ability of Mtb bacteria to oxidize thioamides and thiocarbonyl drugs. Given this predictability, Applicants maintain that the nine species of mutations shown in the specification support a genus claim to mutations in the EtaA gene.

Factor 8: The quantity of experimentation needed to make or use the invention based on the content of the disclosure

The eighth, and last, factor is the amount of experimentation necessary for an artisan to make or use the invention as claimed. The test here is not merely quantitative, since extended experimentation may be permissible if it is routine and if the specification gives sufficient guidance. MPEP § 2164.06, citing *Wands*.

In the instant case, the EtaA gene encodes a protein of only 489 amino acids. An Mtb with any given frameshift, missense or nonsense mutation, or amino acid substitution in the encoded protein can be readily tested for its activity in oxidizing thioamides or thiocarbonyls by assaying for monooxygenase functionality, as taught in the specification at page 17, line 18, to page 18, line 2, or by detecting the presence of the metabolic product, as taught at page 18, lines 3-20.

Thus, only modest experimentation would be required to practice the invention, and clear guidance is given by the specification on how to do it.

Thus, the specification sets forth all the information an artisan needs to test whether any particular protein falls within the "metes and bounds of the claims.

g. The rejection of claim 16

Applicants observe that, in any event, the rejections under §112, 1st paragraph, would apply with much less force to claim 16, which claimed not all thioamides and thiocarbonyls, but the specific drugs ethionamide, thiacetazone, and thiocarlide, the exemplar thioamides and thiocarbonyl whose oxidation the mutations were shown in the specification to affect. Claim 16 has been cancelled and its recitations incorporated into new claims 34-48.

D. Rejection of Claims 12 and 29 As Indefinite

The Action rejects claims 12 and 29 under 35 U.S.C. § 112, second paragraph, as indefinite. In this regard, the Action points out that the preamble of the claims recites the gene, whereas the recited group comprises mutations in the gene. The claims have been amended to correct the correspondence between the preamble and the body of the claim, as helpfully suggested by the Examiner.

Applicants respectfully submit that the claims as amended are definite. Reconsideration and withdrawal of the rejection are respectfully requested.

E. Rejection of Claim 25 As Anticipated

Claim 25 is rejected under 35 U.S.C. §102 (b) as anticipated by the Boehringer Mannheim 1997 Biochemicals catalog. According to the Action, the catalog teaches a container filled with a mixture of all possible hexamer nucleotides for random-primed DNA.

The hexamers of the catalog would non-specifically bind throughout a genome and result in the promiscuous amplification of portions of genes and of intergenic regions. Claim 25 has been amended to recite that the kit contains primers for specifically amplifying the EtaA gene. The non-specific amplification of genome permitted by the hexamers of the catalog therefore do not meet the claim recitations. As amended, the claim is therefore free of the rejection.

F. The Rejection of Claim 25 as Obvious

Claim 25 is rejected under 35 U.S.C. §103 (a) as obvious over an unpublished sequence deposited by Badcock and Churcher (hereafter, "Badcock"), in view of Philipp et al., PNAS USA 93:3132-3137 (1996) ("Philipp") and further in view of Ahern, The Scientist, 1995 ("Ahern"). Applicants traverse.

Badcock sets forth certain sequences from the genome of Mtb, and denotes that nucleotides 14983 to 16452 encode a probable monooxygenase. Philipp is a report of an integrated map of the Mtb genome. Ahern is simply an article on the advantages of kits. According the Action, Badcock teaches the EtaA gene of Mtb and its possible function as a monooxygenase. Action, at page 10. Philipp is stated to establish an ordered set of DNA fragments and supposedly, to teach using primers specific to the Mtb sequence of genomic DNA to facilitate gene mapping.

As noted, the rejection under §103(a) relies on combining Badcock, Philipp and Ahern. A combination of references requires the Action to show some teaching, suggestion, or motivation in the art to combine the references. *See, e.g., In re Geiger*, 815 F.2d 686 (Fed. Cir. 1987). As the Federal Circuit has noted, "[c]ombining prior art references without evidence of such a suggestion, teaching, or motivation simply takes the inventor's disclosure as a blueprint for piecing together the prior art to defeat patentability--the essence of hindsight. *See, e.g., Interconnect Planning Corp. v. Feil,* 774 F.2d 1132, 1138, 227 USPQ 543, 547 (Fed. Cir. 1985) ("The invention must be viewed not with the blueprint drawn by the inventor, but in the state of the art that existed

at the time.")." *In re Dembiczak*, No. 98-1498 (Fed. Cir., April 28, 1999). In this case, the Action impermissibly uses the Applicants' disclosure as the blueprint for combining the referenced patents.

The Federal Circuit Court of Appeals has called the requirement to show a teaching, suggestion, or motivation a "**critical** safeguard against hindsight analysis and rote application of the legal test for obviousness." *In re Rouffet*, 47 USPQ2d 1453 (1998) (emphasis added). The Court explained: "To prevent the use of hindsight based on the invention to defeat patentability of the invention, this court requires the examiner to show a motivation to combine the references that create the case of obviousness." *Id.* The Court elaborated that: "[i]n other words, the examiner must show reasons that the skilled artisan, confronted with the same problems as the inventor and no knowledge of the claimed invention, **would select the elements from the cited prior art references for combination in the manner cited.**" *Id.* (Emphasis added.)

The present Action is based on a hindsight reading of the references. Badcock is merely a deposit of over 38,000 nucleotide bases of the Mtb genome, with a number of coding regions designated as having known or unknown possible functions. The gene that Badcock identifies is of course not designated as EtaA, for that is the designation the Applicants assigned after determining the function of the encoded protein (see, specification at page 7, paragraph 20 and page 8, paragraph 21). The designation for Badcock's putative gene was accession "MTCY01A6.14," and it does identify the gene as a probable monooxygenase. As noted in the specification, at page 7, lines 30-32, "When the tuberculosis genome was sequenced and analyzed in 1998, the gene was considered to bear homology to a bacterial monooxygenase, but was sufficiently different to be classified as a separate, unknown family. Moreover, its substrate was unknown." There is therefore no motivation apparent in Badcock to combine this particular gene, of undetermined function and unknown substrate, with Philipp or Ahern. The only reason to focus on the gene eventually renamed as EtaA is the teaching of the present specification.

Philipp does not supply the motivation lacking in Badcock. Once the genome map was created, and the ordered DNA fragments were created, there was no reason remaining other than possible academic interest to amplify the section containing accession MTCY01A6.14. It is true that Ahern teaches that kits are cost effective. What Ahern teaches, though, is buying pre-made kits for something one is actually going to use. It provides no reason why one of skill in the art would be motivated to create a kit for amplifying a gene that has a protein with an undefined function and unknown substrate -- unless it is read with knowledge of the present disclosure.

This motivation cannot exist simply because there is a possible research interest in determining the gene's function. The Action's position proves too much. The human genome has now been sequenced and made publicly available. Many of the genes are as yet of unknown function. But, if the Action's position were correct, there would be motivation to have kits of primers for all these unknown genes. If one of these genes is later determined next year to be a definitive marker for cancer, under the Action's position, claims to kits to primers for specifically amplifying that gene would not be available even though no function was known when the gene was deposited.

In short, the Action fails to show, in the words of *Rouffet*, "reasons that the skilled artisan, confronted with the same problems as the inventor and no knowledge of the claimed invention, would select the elements from the cited prior art references for combination in the manner cited." In the absence of such a showing, the conclusion is warranted that the Action has simply picked and chosen from the art to recreate the invention. Applicants therefore respectfully submit that the Examiner has failed to present a proper prima facie case of obviousness. The rejection should be reconsidered and, upon reconsideration, withdrawn.

CONCLUSION

In view of the foregoing, Applicants believe all claims now pending in this Application are in condition for allowance. The issuance of a formal Notice of Allowance at an early date is respectfully requested.

If the Examiner believes a telephone conference would expedite prosecution of this application, please telephone the undersigned at 415-576-0200.

Respectfully submitted,

Laurence J. Hyman
Reg. No. 35,551

TOWNSEND and TOWNSEND and CREW LLP
Two Embarcadero Center, 8th Floor
San Francisco, California 94111-3834
Tel: 415-576-0200
Fax: 415-576-0300
LJH:ljh

Enclosure: Sreevatsan

VERSION WITH MARKINGS TO SHOW CHANGES MADE

IN THE SPECIFICATION:

[30] References here to "MTb" refer to *Mycobacterium tuberculosis*. The sequence of the entire genome of MTb is set forth in TubercuList, which can be found [at] on the Internet by typing "http://" followed by "genolist.pasteur.fr/TubercuList/."

[62] In a particularly preferred embodiment, the EtaA gene can be amplified using the primers 5'-GGGGTACCGACATTACGTTGATAGCGTGGA-3' (SEQ ID NO:3) and 5'-ATAAGAATGCGGCCGCAACCGTCGCTAAAGCTAAACC-3' (SEQ ID NO:4) (EtaA). Many other primer sets can be selected using standard programs widely available in the art. For example, the program "Primer3" [is available] can be found on-line [at] by typing "www-" followed by "genome.wi.mit.edu/cgi-bin/primer/primer3_www.cgi." This program was used to select the primer pairs noted above, using the default conditions. The program was also used to select the following sequencing primers, which can be used to amplify sections of the EtaA gene for sequencing:

5' ATCATCCATCCGCAGCAC 3' (SEQ ID NO:5);
5' AAGCTGCAGGTTCAACC 3' (SEQ ID NO:6);
5' GCATCGTGACGTGCTTG 3' (SEQ ID NO:7);
5' AAGCTGCAGGTTCAACC 3' (SEQ ID NO:8);
5' TGAACTCAGGTCGCGAAC 3' (SEQ ID NO:9);
5' AACATCGTCGTGATCGG 3' (SEQ ID NO:10);
5' ATTTGTTCCGTTATCCC 3' (SEQ ID NO:11);
5' AACCTAGCGTGTACATG 3' (SEQ ID NO:12);
5' TCTATTTCCCATCCAAG 3 (SEQ ID NO:13); and

PATENT

5' GCCATGTCGGCTTGATTG 3' (SEQ ID NO:14).

[81] The production of metabolite (**5**) from ETA by tuberculosis is surprising as 4-pyridylmethanol is a major metabolite of INH by whole cells of MTb (Youatt, J. *Aust J Chem* 14:308 (1961); Youatt, J. *Aust J Exp Biol Med Sci* 38:245 (1960); Youatt, J. *Aust J Biol Med Sci* 40:191 (1962)). Like spontaneous oxidation of INH, spontaneous oxidation of ETA fails to produce any trace of the major *in vivo* metabolite, (2-ethyl-pyridin-4-yl)methanol. INH has been shown to be activated by KatG *in vitro* to a variety of products including isonicotinic acid, isonicotinamide and isonicotinaldehyde (which *in vivo* is rapidly reduced to 4-pyridylmethanol) (Johnsson, K. et al., *J Am Chem Soc* 116:7425 (1994)). INH metabolism to 4-pyridylmethanol only occurs in drug-susceptible organisms while drug-resistant organisms no longer produce this metabolite (Youatt, J., *Am Rev Respir Dis* 99:729 (1969)). Similarly, we postulate that ETA is activated via the corresponding S-oxide to a sulfinate that can form an analogous aldehyde equivalent (an imine) through a radical intermediate (Paez, O.A. et al., *J Org Chem* 53:2166 (1988)). [(Figure 5)]

[88] INH (**6**) has been shown to be activated by KatG in vitro to a variety of products including isonicotinic acid, isonicotinamide and isonicotinaldehyde (**9**) (which *in vivo* is rapidly reduced to 4-pyridylmethanol (**10**)) (Johnsson, K. & Schultz, P. G., J Am Chem Soc 116:7425-68 (1994)). The results support the notion that in vivo INH is metabolized by oxidation to an acyl diimide (**7**), then to a diazonium ion (**8**) or an isonicotinyl radical which may abstract a hydrogen atom from a suitable donor to form isonicotinaldehyde. Similarly, we postulate that ETA is activated via the corresponding S-oxide (**2**) to a sulfinate that can form an analogous aldehyde equivalent (an imine) through a radical intermediate [(Figure 5)]. Hydrolysis of this imine could be followed by reduction of the resulting aldehyde to the observed metabolite (**5**).

IN THE CLAIMS:

1. (Amended) A method of determining the ability of a *Mycobacterium tuberculosis* bacterium to oxidize a thioamide or a thiocarbonyl, said method comprising detecting a mutation in an EtaA gene (SEQ ID NO:1) in said bacterium, which mutated gene encodes an amino acid sequence which differs from that of SEQ ID NO:2, wherein detection of the mutation is indicative of decreased ability to oxidize a thioamide or a thiocarbonyl.

12. (Amended) A method of claim 11, wherein said [known] mutation in said EtaA gene is selected from the group consisting of (a) a frameshift mutation consisting of a deletion at position 65, an addition at position 567, and an addition at position 811, and (b) a single nucleotide polymorphism which causes an amino acid substitution selected from the group consisting of: G43C, P51L, D58A, Y84D, T186K, T342K, and A381P.

21. (Amended) A method of screening an individual for a *Mycobacterium tuberculosis* bacterium resistant to treatment by a thioamide or a thiocarbonyl drug, comprising

(a) obtaining a biological sample containing said bacterium from said individual, and

(b) detecting a mutation in an EtaA gene (SEQ ID NO:1) in said bacterium, which mutated gene encodes an amino acid sequence which differs from that of SEQ ID NO:2, wherein detection of the mutation is indicative said bacterium is resistant to treatment by a thioamide or a thiocarbonyl drug.

25. (Amended) A kit for determining the ability of a *Mycobacterium tuberculosis* bacterium to oxidize a thioamide or a thiocarbonyl, the kit comprising:

(a) a container, and

(b) primers for specifically amplifying an EtaA gene of said bacterium or a portion of said EtaA gene containing a mutation affecting the ability of the bacterium to oxidize a thioamide.

29. (Amended) A kit of claim 28, wherein said mutation in said [mutated] EtaA gene is selected from the group consisting of (a) a frameshift mutation consisting of a deletion at position 65, an addition at position 567, and an addition at position 811, and (b) a single nucleotide polymorphism which causes an amino acid substitution selected from the group consisting of: G43C, P51L, D58A, Y84D, T186K, T342K, and A381P

Proc. Natl. Acad. Sci. USA
Vol. 94, pp. 9869–9874, September 1997
Microbiology

Restricted structural gene polymorphism in the *Mycobacterium tuberculosis* complex indicates evolutionarily recent global dissemination

SRINAND SREEVATSAN*, XI PAN*, KATHRYN E. STOCKBAUER*, NANCY D. CONNELL†, BARRY N. KREISWIRTH‡, THOMAS S. WHITTAM§, AND JAMES M. MUSSER*¶

*Section of Molecular Pathobiology, Department of Pathology, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030; †Department of Microbiology and Molecular Genetics, and National Tuberculosis Center, New Jersey Medical School, Newark, NJ 07103; ‡Tuberculosis Center, Public Health Research Institute, 455 First Avenue, New York, NY 10016; and §Institute of Molecular Evolutionary Genetics, Department of Biology, Mueller Laboratory, Pennsylvania State University, University Park, PA 16802

Communicated by Barry R. Bloom, Albert Einstein College of Medicine, Hastings-on-Hudson, NY, July 11, 1997 (received for review April 11, 1997)

ABSTRACT **One-third of humans are infected with *Mycobacterium tuberculosis*, the causative agent of tuberculosis. Sequence analysis of two megabases in 26 structural genes or loci in strains recovered globally discovered a striking reduction of silent nucleotide substitutions compared with other human bacterial pathogens. The lack of neutral mutations in structural genes indicates that *M. tuberculosis* is evolutionarily young and has recently spread globally. Species diversity is largely caused by rapidly evolving insertion sequences, which means that mobile element movement is a fundamental process generating genomic variation in this pathogen. Three genetic groups of *M. tuberculosis* were identified based on two polymorphisms that occur at high frequency in the genes encoding catalase-peroxidase and the A subunit of gyrase. Group 1 organisms are evolutionarily old and allied with *M. bovis*, the cause of bovine tuberculosis. A subset of several distinct insertion sequence IS*6110* subtypes of this genetic group have IS*6110* integrated at the identical chromosomal insertion site, located between *dnaA* and *dnaN* in the region containing the origin of replication. Remarkably, study of ≈6,000 isolates from patients in Houston and the New York City area discovered that 47 of 48 relatively large case clusters were caused by genotypic group 1 and 2 but not group 3 organisms. The observation that the newly emergent group 3 organisms are associated with sporadic rather than clustered cases suggests that the pathogen is evolving toward a state of reduced transmissability or virulence.**

One-third of the world's population is infected with *Mycobacterium tuberculosis*, and 3 million human deaths annually are attributed to the organism (1, 2). Although there is a very large global pool of infected individuals and considerable chromosomal heterogeneity based on restriction fragment length polymorphism (RFLP) patterns generated by probing with mobile insertion elements (3, 4), studies of drug resistance and pathogenesis have raised the possibility that synonymous (silent) nucleotide substitutions in structural genes may be limited (5). To investigate this apparent discrepancy from the perspective of molecular population genetics, we sequenced two megabases in 26 structural genes or loci in strains of *M. tuberculosis* and the three closely related members of the *M. tuberculosis* complex (*M. africanum*, *M. bovis*, and *M. microti*) collected worldwide.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

MATERIALS AND METHODS

Bacterial Isolates. The study is based on a sample of 842 *M. tuberculosis* complex isolates recovered from diverse geographic localities. The organisms include *M. tuberculosis* (n = 715), *M. bovis* (n = 109), and *M. africanum* and *M. microti* (n = 9 each). *M. tuberculosis* isolates were recovered from diseased patients in the United States (five states), Latin America (Mexico, Honduras, Ecuador, Peru, Venezuela, Brazil, and Chile), Europe (Portugal, Spain, The Netherlands, Belgium, Germany, Switzerland, Italy, former Yugoslavia, and Romania), Africa (Kenya, Rwanda, Guinea, Algeria, Somalia, and Zaire), the mid-East (Yemen, Israel, Turkey, and Iran), and elsewhere (Australia, Burundi, China, India, Japan, South Korea, Mongolia, Nepal, Philippines, Sri Lanka, Tahiti, Thailand, and Vietnam) (6). This collection of *M. tuberculosis* isolates represents the range of insertion sequence IS*6110* fingerprint diversity in the species (refs. 3 and 4 and unpublished data) and includes organisms recovered from patients with pulmonary and extrapulmonary tuberculosis. Organisms in the sample had from 0 to 21 copies of IS*6110*. Moreover, organisms classified into several groups based on a multiplex PCR analysis (7) were included in the analysis.

The 109 *M. bovis* isolates were recovered in five countries and from eight host species (6). The *M. microti* specimens were recovered from voles (n = 7), a pig (n = 1) in the Netherlands, and a rock hyrax (n = 1) in South Africa. Isolates of *M. africanum* were recovered from patients with tuberculosis living in Sierra Leone, Africa.

Species assignment of isolates of *M. tuberculosis, M. africanum, M. bovis,* and *M. microti* was based on analysis of accepted phenotypic criteria.

IS*6110* RFLP Profiling. IS*6110* RFLP profiling was performed by an internationally standardized method with restriction endonuclease *Pvu*II (3). Hybridizing bands were visualized by enhanced chemiluminescence, and banding patterns were compared with computer-assisted image analysis.

PCR Amplification and Gene Sequencing. All or part of 26 genes were characterized (Table 1). Oligonucleotide primers used to amplify the target regions are available by request from J.M.M. Sequence data were generated with an Applied Biosystems model 373A or 377 instrument. DNA sequence data reported previously (8) were also included in this analysis. A total of 2,150,267 bp was characterized.

Determination of the Presence of IS*6110* Between *dnaA* and *dnaN*. A total of 162 strains were studied by PCR for the presence of IS*6110* between *dnaA* and *dnaN*. We used PCR

Abbreviations: RFLP, restriction fragment length polymorphism; IS, insertion sequence.

¶To whom reprint requests should be addressed. e-mail: jmusser@path.bcm.tmc.edu.

Table 1. Regions of *M. tuberculosis* complex genes analyzed for nucleotide sequence diversity

Gene	Size, bp	Product	No. of strains analyzed	Nucleotides sequenced, positions	No. of sites with silent variation	GenBank accession no.
katG	2,170	Catalase-peroxidase	55	1,991–4,160	5	X68081
katG	570	Catalase-peroxidase	360	2,881–3,330	2	X68081
rpoB	350	RNA polymerase, beta subunit	305	83–432	2	L05910
mabA	700	Fatty acid biosynthesis enzyme	46	1–700	1	U02492
mabA	468	Fatty acid biosynthesis enzyme	92	65–533	1	U02492
inhA	805	Enoyl reductase	72	986–1,794	0	U02492
gyrA	318	DNA gyrase, A subunit	629	2,384–2,701	6	L27512
gyrB	351	DNA gyrase, B subunit	17	1,580–1,930	0	L27512
hsp65	366	65-kDa heat shock protein	267	453–818	0	M15467
rpsL	350	Ribosomal protein S12	178	10–359	1	L08011
rrs	1,042	16S rRNA	122	1–1,046	0	X52917
aroA	349	5-Enolpyruvylshikimate-3-P synthase	8	714–1,063	0	M62708
recA	349	RecA protein	9	1,646–1,995	0	X58485
ahpC	588	Alkylhydroperoxide reductase	97	628–1,216	0	U16243
ahpC (upstream)	421	Alkylhydroperoxide reductase	196	981–1,401	0	U16243
*oxyR**	528	Pseudogene	117	1–528	5	U16243
rpoV	288	Principal sigma factor	48	2,192–2,480	0	U21134
16-kDa Ag	656	16-kDa antigen	35	25–680	0	S79751
embC-A	568	Glycosyltransferase	101	80–647	0	U68480
embA	1,197	Glycosyltransferase	43	2,720–3,916	1	U68480
embA	1,850	Glycosyltransferase	7	920–2,770	0	U68480
embB	3,295	Glycosyltransferase	19	1–3,294	1	U68480
embB	2,560	Glycosyltransferase	3	640–3,200	0	U68480
embB	1,952	Glycosyltransferase	86	98–2,050	0	U68480
mpcA	1,563	Phospholipase	30	438–2,000	2	U49511
mpcA–B	501	Phospholipase region	35	1,748–2,249	0	U49511
pncA	488	Pyrazinamidase	50	30–518	0	U59967
pncA	630	Pyrazinamidase	131	−82 to 558	2	U59967
Monooxygenase (upstream)	340		29	11,121–11,460	0	Z80108
pzaA	1,567	Pyrazinamidase	30	54–1,620	0	†
mdh	1,551	Malate dehydrogenase	34	928–2,478	0	†
ndh	1,597	NADH dehydrogenase	56	921–2,517	0	†
ideR	850	DtxR homolog	22	−120 to 730	1	U14191

Ag, antigen.
**oxyR* is a pseudogene and, therefore, all polymorphisms are silent.
†Unpublished sequence.

analysis with the following oligonucleotide primers: forward, 5′-TCCGAGATGGCCGAGCGCCG-3′; reverse, 5′-CCACCCACGACACCGCATCG-3′. Strains with the insert yielded an amplicon of ≈1,800 bp, and those without the insert had an ≈700-bp PCR product.

Assignment of Strains to Three Genotypic Groups. Isolates of *M. tuberculosis* complex organisms were assigned to one of three genotypic groups based on the combinations of polymorphisms at *katG* codon 463 and *gyrA* codon 95. Group 1 has the allele combination *katG* codon 463 CTG (Leu) and *gyrA* codon 95 ACC (Thr); group 2 has *katG* 463 CGG (Arg) and *gyrA* codon 95 ACC (Thr), and group 3 organisms have *katG* 463 CGG (Arg) and *gyrA* codon 95 AGC (Ser). Polymorphism located at *katG* codon 463 was identified by automated DNA sequencing (9) or a PCR–RFLP strategy with restriction endonuclease *Nci*I or *Msp*I (10). Polymorphism occurring at *gyrA* codon 95 was indexed by automated DNA sequencing (11). Strains (n = 850) recovered from patients in Houston from June 1995 to February 1997 were assigned to one of the three genotypic groups. Note that only approximately 100 strains of this group of 850 were included in the sample of 842 organisms used for sequencing. Case clusters were defined as groups of five or more patients infected with the same *M. tuberculosis* strain, based on IS*6110* pattern analysis and detailed epidemiologic contact investigation data. Because the New York City strain repository contains greater than 5,000 organisms recovered between 1991 and 1997, we used a two-pronged sampling strategy. First, random isolates representing each of the 17 IS*6110* types with 20 or more organisms were analyzed. Case clusters with 20 or more patients were studied because with ≈5,000 strains of *M. tuberculosis* in the database, it would be a considerable undertaking to study all clusters with five or more members. Next, a random sample of 75 isolates with unique IS*6110* typing patterns was selected for analysis. We also analyzed a sample of 25 isolates recovered from a statewide survey of *M. tuberculosis* cases in New Jersey between June 1995 and December 1996. Approximately 300 isolates were obtained during this period, and about 250 of these organisms had unique IS*6110* types. The 25 isolates analyzed represented a random sample of 21 strains with unique IS*6110* profiles and four organisms causing case clusters. New Jersey case clusters were defined as groups of five or more patients infected with *M. tuberculosis* having the same IS*6110* type. The New Jersey organisms were not included in the sample of 841 strains used for the nucleotide sequence database.

RESULTS

Allelic Variation Is Largely Associated with Antibiotic Resistance. Compilation of the two megabases of sequence data for the 26 genes revealed that greater than 95% of nucleotide substitutions caused amino acid replacements or other mutations in gene regions linked to antibiotic resistance

Table 2. Levels of allelic polymorphism recorded in genes of pathogenic bacteria

Organism	No. of genes or gene segments	D_s*	Relative variation†
M. tuberculosis complex	26	<0.01	1
Shigella sonnei	2	0.01	1
Vibrio cholerae	1	0.41	41
Streptococcus pyogenes‡	3	1.02	102
Neisseria meningitidis	4	6.18	618
M. avium-intracellulare	1	10.10	1,010
Escherichia coli	11	11.77	1,177
Borrelia burgdorferi	3	≈20	≈2,000
Salmonella enterica	5	42.08	4,208

*Average number of synonymous substitutions per 100 synonymous sites (28), calculated on the basis of sequence information published elsewhere (12-27). The value reported for *S. sonnei* includes both synonymous and nonsynonymous sites and is based on RFLP analysis (15).
†Calculated relative to the D_s value of the *M. tuberculosis* complex.
‡Includes some unpublished data.

and driven to high frequency by direct drug selection (5). Over all 26 genes examined, only 32 polymorphic nucleotide sites were identified that have not been directly associated with antibiotic resistance. Of these, 30 were synonymous (silent) changes. The 30 substitutions occurred in *katG* (n = 7), *gyrA* (n = 6), *oxyR* (n = 5), *rpoB*, *mabA*, *pncA*, and *mpcA* (n = 2 each), and *embA*, *embB*, *ideR*, and *rpsL* (n = 1 each). Although there was substantial chromosomal heterogeneity in RFLP patterns generated by probing with mobile elements, the striking lack of silent substitutions in *M. tuberculosis* complex members from global sources is unexpected given that 1 billion humans carry the pathogen and, hence, the bacterial population size worldwide must be enormous.

Allelic Variation in *M. tuberculosis* Complex Organisms Is Restricted Compared with Other Pathogenic Bacteria. The level of silent nucleotide variation in *M. tuberculosis* complex members is greatly restricted and considerably less than observed in other pathogenic bacteria (12–27) (Table 2). Importantly, the organisms studied include other mycobacteria, and species such as *Neisseria meningitidis* that are strict host specialists. Thus, although members of the *M. tuberculosis* complex have preferred hosts, ecological host specialization *per se* does not account for the restricted genetic variation observed in the *M. tuberculosis* complex. To put the level of allelic variation in *M. tuberculosis* in perspective, it is even less than that in *Shigella sonnei*, which genetically is merely a specialized distinct pathogenic clone of *Escherichia coli* (29).

Identification of Three Genotypic Groups. Inspection of the sequence data revealed that only the variants at *katG* codon 463 and *gyrA* codon 95 were present at high frequency. These two sites apparently do not participate in antibiotic resistance (30, 31) and, hence, were used as genetic markers that record the history of organism divergence. All members of the *M. tuberculosis* complex were assigned to one of three distinct genotypic groups based on the combination of polymorphisms located at these two sites (Fig. 1). All isolates of *M. bovis*, *M. microti*, and *M. africanum* studied had the combination of polymorphisms characteristic of genotypic group 1. In contrast, *M. tuberculosis* isolates fell into each of the three groups.

IS*6110* Frequency Distribution Among Isolates of the Three Genotypic Groups. To determine the extent to which organisms assigned to the three genotypic groups have diverged in other properties, we examined the distribution of IS*6110*, an insertion sequence present in most *M. tuberculosis* isolates. The analysis was based on study of 421 organisms recovered from patients in Houston, TX, in a 13-month period. To avoid bias that would be generated by including epidemiologically associated organisms, only one isolate was included from each subtype identified by standard IS*6110* typing. The frequency distribution of IS*6110* copy number differed among the three genotypic groups (Fig. 2). Moreover, there was little sharing of related IS*6110* profiles among groups. These data indicate that isolates of the three groups have accumulated differences in IS*6110* copy number and pattern and undergone rapid chromosomal differentiation relative to silent mutations.

Correlation of Site of IS*6110* Insertion and Genotypic Group. Two distinct molecular mechanisms could account for the occurrence of three genotypic groups of *M. tuberculosis*. One mechanism postulates that organisms assigned to each of the three genotypic groups have shared a common ancestor, and the extensive IS*6110* variation recorded among members of each group arose after the divergence of the last common ancestor for each group. An alternative idea postulates that

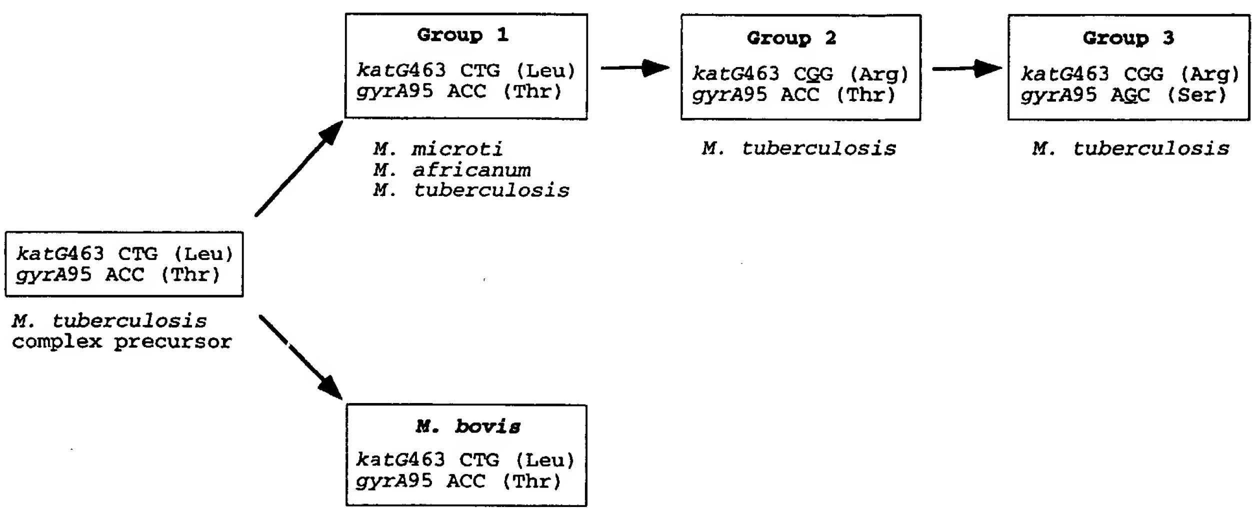

FIG. 1. Broad evolutionary scenario for *M. tuberculosis* complex organisms. The precursor of *M. tuberculosis* complex organisms was characterized by KatG codon 463 (Leu) and GyrA codon 95 (Thr). Strain Ravenal is a typical *M. bovis* isolate; New York City IS*6110* type strain W and Houston IS*6110* types 002, 003, 007, 015, and 033 are group 1 organisms; Erdman, Oshkosh, New York City strain C, and Houston IS*6110* types 004, 006, 016, 020, and 030 are group 2; and H37Ra, H37Rv, and Houston IS*6110* type 001 are group 3.

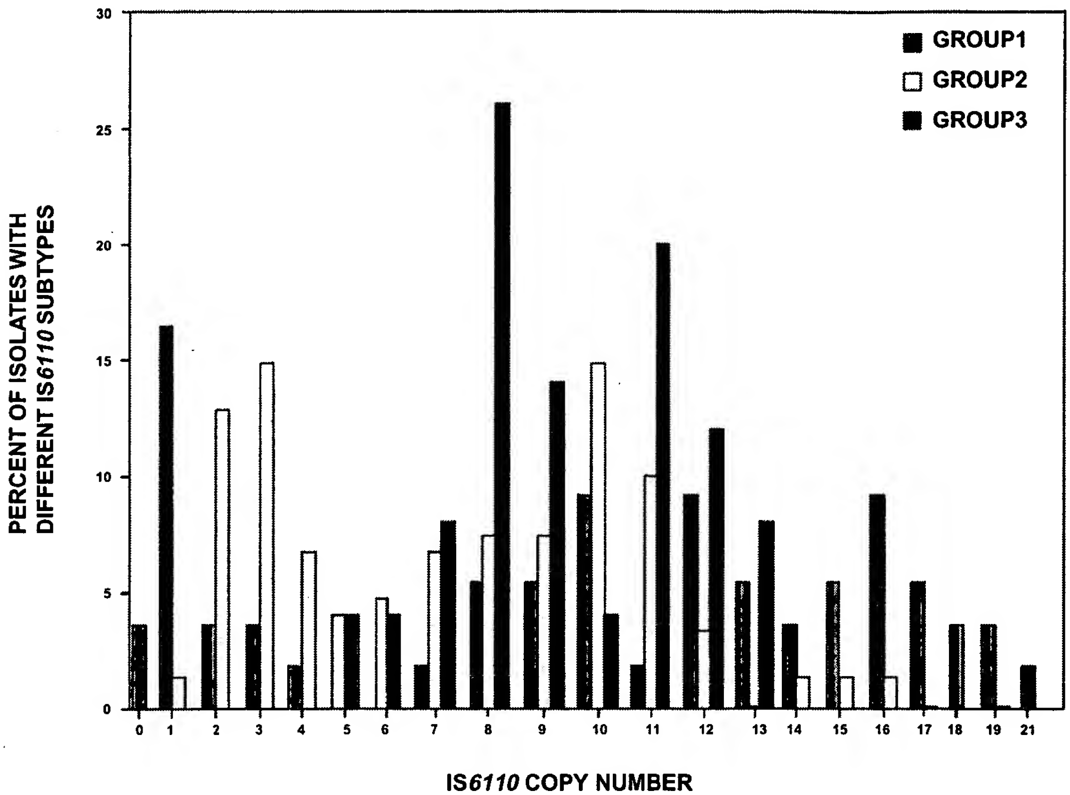

FIG. 2. Distribution of IS*6110* copy number with respect to three genotypic groups of *M. tuberculosis*. The data are based on analysis of 427 isolates recovered from patients in Houston, TX. To avoid bias caused by redundant sampling of recent derivatives of epidemiologically linked isolates, only one isolate per distinct IS*6110* subtype was used in the analysis. This resulted in a data set composed of 213 distinct IS*6110* subtypes, of which 21% were group 1, 62% group 2, and 18% group 3.

organisms have repetitively converged by independent evolutionary pathways to form the three genotypic groups. The distribution of IS*6110* copy number described above is strong evidence against the convergence hypothesis. To differentiate more fully between the two hypothesis, we used PCR to examine the distribution of one randomly chosen IS*6110* insertion site located between *dnaA* and *dnaN* among the three groups marked by *gyrA* and *katG* polymorphisms. IS*6110* was located at the target site in 26% of 97 organisms in group 1 having distinct IS*6110* profiles. In contrast, all 65 randomly chosen isolates of groups 2 and 3 tested lacked the element at this site. All 32 isolates of *M. bovis*, *M. africanum*, and *M. microti* studied also lacked the element at this site. The confinement of this insertion site polymorphism to a subset of group 1 *M. tuberculosis* further demonstrates an historical separation and genotypic differentiation of the three groups.

Additional documentation of a common ancestry of many group 1 isolates with IS*6110* located between *dnaA* and *dnaN* was sought by DNA sequencing. Analysis of six randomly chosen strains with different IS*6110* patterns found that all six organisms had IS*6110* integrated at the identical site located 66 bp downstream from *dnaA*. The integration of IS*6110* at precisely the same position in these group 1 members adds further strong support to the idea that the organisms have shared a common ancestor.

Genotypic Group 1 Organisms Are Ancestral to *M. tuberculosis* Groups 2 and 3. In the proposed evolutionary pathway presented in Fig. 1, genotypic group 1 organisms are especially important because they link the predominantly nonhuman pathogens *M. microti* and *M. bovis* and the human specialists *M. africanum* and *M. tuberculosis*. If our evolutionary hypothesis is correct, then we expect that group 1 organisms carry a greater level of genetic diversity than group 2 and group 3 bacteria because they have had a longer time to evolve and thereby accumulate variation. Three facts support this hypothesis. (*i*) Group 1 organisms have the broadest range of IS*6110* copy number, with 0–21 copies of this element. (*ii*) Group 1 organisms had 73% of all synonymous substitutions identified, despite being only about 15% of all organisms analyzed. (*iii*) All five variant sites detected in the pseudogene *oxyR* were either in *M. bovis* or group 1 organisms, and none were present in organisms assigned to groups 2 or 3. The neutral theory of molecular evolution predicts that allelic variation in pseudogenes exceeds that present in genes encoding protein products because functional constraints are relaxed (32). The observation of increased allelic variation in *oxyR* in group 1 organisms provides critical independent support of the concept that group 1 organisms are ancestral to groups 2 and 3. Also consistent with our thesis that group 1 bacteria are ancestral to groups 2 and 3 is the fact that all four organisms in our collection that lack IS*6110* are group 1 members. Absence of IS*6110* is the expected condition for a primitive organism from which *M. tuberculosis* containing this element may have arisen.

DISCUSSION

Restricted Allelic Variation. The remarkably restricted variation in *M. tuberculosis* complex structural genes has important ramifications for studies of virulence and drug resistance. The lack of allelic diversity means that when amino acid polymorphisms, or regulatory region nucleotide variation are observed, there should be strong suspicion that the variation has functional consequences, such as antibiotic resistance (5). Restricted allelic diversity also means that identification of pro-

teins with immunoprophylaxis or diagnostic utility would be of significant clinical value due to the very low probability for variation globally. Conversely, should future genomic analyses uncover highly polymorphic genes, the finding could indicate that the diversity is being driven by host immunologic responses.

The spontaneous mutation frequency of *M. tuberculosis* is in the range recorded for most other bacteria (33), an observation that permits us to exclude the possibility that an unusual DNA repair phenotype or replication fidelity is responsible for the lack of accumulation of neutral variation. A small population size also is unlikely to explain the limited allelic diversity given that one-third of the world's population is infected with this bacterium. We believe that the only reasonable hypothesis to account for the lack of allelic variation is that *M. tuberculosis* underwent a recent evolutionary bottleneck, presumably at the time of speciation estimated to have occurred roughly 15,000 to 20,000 years ago (8).

Two main processes are driving a considerable proportion of genomic differentiation in *M. tuberculosis*. One mechanism responsible for creating variation is transposition of IS*6110* and other mobile elements, and these events can rapidly generate new subclones (34). For example, a single multidrug-resistant organism marked by a distinct IS*6110* pattern and known as strain W within years spawned at least seven variants differing by one or two copies of IS*6110* (34). The level of species variation also is being profoundly increased by humans through antibiotic use. In our two-megabase data set, greater than 95% of nucleotide changes were directly associated with antibiotic resistance. All antibiotic resistance in *M. tuberculosis* is chromosomally mediated, unlike most pathogenic bacteria in which drug resistance is mediated by plasmid-borne genes that can be discarded when selection pressure is removed. This means that the genomic changes arising from drug selection will be removed from the species gene pool only if the resistant organisms themselves are extinguished. Given the considerable difficulties in successfully treating drug-resistant strains, and the lack of agents to eradicate latent infections, most of the resistance-associated diversity is likely to be maintained in the *M. tuberculosis* population.

Genotypic Group 1 and Group 2 Organisms Are Disproportionately Represented Among Clustered Cases of Tuberculosis. One theme that has emerged from the study of bacterial population genetics is that there are frequently important biomedical and ecological correlates of population structure. Stated another way, groups of related bacterial genotypes tend to behave nonrandomly. This is especially true for species (like *M. tuberculosis*) in which horizontal gene transfer contributes little to the generation of genomic diversity.

The availability of extensive databases containing IS*6110* typing data for ≈6,000 isolates permitted us to determine whether organisms of the three groups were equally represented among genotypes causing case clusters of tuberculosis. The Houston *M. tuberculosis* database has typing information for all 850 strains recovered over a 2-year period in a comprehensive population-based survey. Detailed epidemiologic data, including contact investigation, are available for all patients from whom these organisms were cultured. Interestingly, 25 of 26 case clusters involving five or more patients were caused by genotypic group 1 and 2 organisms (Table 3). This striking association was confirmed by analysis of strains recovered in metropolitan New York City. All 21 case clusters in the New York City area were also caused by genotypic group 1 and 2 organisms. Nonrandom association of group 1 and group 2 organisms with clustered cases was independent of the drug susceptibility phenotype of the strains. Critically, the dominance of group 1 and group 2 organisms as causes of case clusters was not due to the formal albeit trivial explanation that group 3 organisms do not occur in these diverse localities (Table 3).

Variance in pathogen behavior is likely to be the result of two processes. Nucleotide changes may result in amino acid alteration in proteins or modification of regulatory sequences that produce increased or decreased expression of protein products. Data have been presented showing that single nucleotide or amino acid changes can significantly alter virulence of *M. tuberculosis* for guinea pigs and virulence in other pathogens (35, 36). A second possibility that has not yet been explored is that the site of chromosomal integration of IS*6110* or other mobile elements participates. Mobile element insertion or excision from genomic sites can alter expression of genes involved in host–pathogen interactions or contribute to bacterial fitness (37, 38). Simonet *et al.* (37) reported that the invasin gene (*inv*) of *Yersinia pestis* (the cause of plague) is inactivated by a 708-bp IS*200*-like element. This single insertion prevents the pathogen from invading host cells, thereby altering the host–pathogen interaction and selecting for new virulence properties. Moreover, the site of IS*1301* insertion in *Neisseria meningitidis* changes expression of cell surface sialic acid, a critical virulence factor in blood and brain infections (38). It is reasonable to anticipate that insights about *M. tuberculosis* pathogenesis will be revealed by systematic study of IS*6110* integration sites and assessment of resulting functional alterations.

Although clearly more work is required to identify the molecular basis for our observation that strains of genotypic group 3 have a reduced capacity to cause large case clusters, we note that the finding has important implications for public health strategies, evolutionary biology, and virtually every aspect of *M. tuberculosis* investigation. Tuberculosis control strategies and transmission models have largely relied on the idea that strains or subclones are roughly equivalent in medically important biological attributes (39). Our data suggest this is not the case. Public health approaches and tuberculosis transmission models may benefit from further refinement by interpreting data in the context of the three genetic groups described herein. Advances in the study of host–pathogen interactions, genome sequences (40), antimicrobial agent resistance (5), and the genetics of host susceptibility (41) are also likely to benefit by the insights into the evolutionary history revealed by our analysis.

Table 3. Distribution of *M. tuberculosis* genotypic groups in Houston and the New York City area

	Genotypic group*					
	1		2		3	
Place	Case clusters, no.	Unique isolates, no.	Case clusters, no.	Unique isolates, no.	Case clusters, no.	Unique isolates, no.
Houston	10 (37%)	100 (28%)	16 (59%)	200 (56%)	1 (4%)†	55 (16%)
New York City area	7 (32%)	23 (23%)	15 (68%)	63 (62%)	0	15 (15%)

*Group 1, allele combination *katG* codon 463 CTG (Leu) and *gyrA* codon 95 ACC (Thr); group 2, *katG* 463 CGG (Arg) and *gyrA* codon 95 ACC (Thr); group 3, *katG* 463 CGG (Arg) and *gyrA* codon 95 AGC (Ser).

†Genotypic group 1 versus group 3, $P \approx 0.008$; genotypic group 2 versus group 3, $P \approx 0.015$; genotypic group 1 versus group 2, not significant. P values were calculated with pooled Houston and New York City area samples.

Note Added in Proof: Analysis of an additional one megabase of sequence data from 10 genes and the DR region confirmed the broad evolutionary scenario described in Fig. 1.

We thank many colleagues who supplied strains and provided critical comments on an early draft of the manuscript. In particular, R. K. Selander contributed scholarly input, and two anonymous reviewers suggested several improvements. G. Adams supplied assistance with statistical analyses. We are especially indebted to K. Davenny and A. M. Ginsberg for their support. This research was supported by Public Health Services Grants DA-09238, AI-37004, and AI-41168 to J.M.M.

1. Bloom, B. R. & Murray, C. J. L. (1992) *Science* **257,** 1055–1064.
2. Raviglione, M. C., Snider, D. E. & Kochi, A. (1995) *J. Am. Med. Assoc.* **273,** 220–226.
3. van Embden, P. D. A., Cave, M. D., Crawford, J. T., Dale, J. W., Eisenach, K. D., Gicquel, B., Hermans, P. W. M., Martin, C., McAdam, R., Shinnick, T. M. & Small, P. M. (1993) *J. Clin. Microbiol.* **31,** 406–409.
4. Hermans, P. W. M., Messadi, F., Guebrexabher, H., van Soolingen, D., de Haas, P. E. W., Heersma, H., de Neeling, H., Ayoub, A., Portaels, F., Frommel, D., Zribi, M. & van Embden, J. D. A. (1995) *J. Infect. Dis.* **171,** 1504–1513.
5. Musser, J. M. (1995) *Clin. M.crobiol. Rev.* **8,** 496–514.
6. Sreevatsan, S., Escalante, P., Pan, X., Gillies, D. A., III, Siddiqui, S., Khalaf, C. N., Kreiswirth, B. N., Bifani, P., Adams, L. G., Ficht, T., Perumaalla, S., Cave, M. D., van Embden, J. D. A. & Musser, J. M. (1996) *J. Clin. Microbiol.* **34,** 2007–2010.
7. Plikaytis, B. B., Marden, J. L., Crawford, J. T., Woodley, C. L., Butler, W. R. & Shinnick, T. M. (1994) *J. Clin. Microbiol.* **32,** 1542–1546.
8. Kapur, V., Whittam, T. S. & Musser, J. M. (1994) *J. Infect. Dis.* **170,** 1348–1349.
9. Musser, J. M., Kapur, V., Williams, D. L., Kreiswirth, B. N., van Soolingen & van Embden, J. D. A. (1996) *J. Infect. Dis.* **173,** 196–202.
10. Cockerill, F. R., III, Uhl, J. R., Temesgen, Z., Zhang, Y., Stockman, L., Roberts, G. D., Williams, D. L. & Kline, B. C. (1995) *J. Infect. Dis.* **171,** 240–245.
11. Kapur, V., Li, L.-L., Hamrick, M. R., Plikaytis, B. B., Shinnick, T. M., Telenti, A., Jacobs, W. R., Jr., Banerjee, A., Cole, S., Yuen, K. Y., Clarridge, J. E., III, Kreiswirth, B. N. & Musser, J. M. (1995) *Arch. Pathol. Lab. Med.* **119,** 131–138.
12. Nelson, K., Whittam, T. S. & Selander, R. K. (1991) *Proc. Natl. Acad. Sci. USA* **88,** 6667–6671.
13. Nelson, K. & Selander, R. K. (1992) *J. Bacteriol.* **174,** 6886–6895.
14. Boyd, E. F., Nelson, K., Wang, F.-S., Whittam, T. S. & Selander, R. K. (1994) *Proc. Natl. Acad. Sci. USA* **91,** 1280–1284.
15. Karaolis, D. K. R., Lan, R. & Reeves, P. R. (1994) *J. Clin. Microbiol.* **32,** 796–802.
16. Karaolis, D. K. R., Lan, R. & Reeves, P. R. (1995) *J. Bacteriol.* **177,** 3191–3198.
17. Li, J., Ochman, H., Groisman, E. A., Boyd, E. F., Solomon, F., Nelson, K. & Selander, R. K. (1995) *Proc. Natl. Acad. Sci. USA* **92,** 7252–7256.
18. Kapur, V., Kanjilal, S., Hamrick, M. R., Li, L.-L., Whittam, T. S., Sawyer, S. A. & Musser, J. M. (1995) *Mol. Microbiol.* **16,** 509–519.
19. Caporale, D. A. & Kocher, T. D. (1994) *Mol. Biol. Evol.* **11,** 51–64.
20. Zhou, J. & Spratt, B. G. (1992) *Mol. Microbiol.* **6,** 2135–2146.
21. DuBose, R. F., Dykhuizen, D. E. & Hartl, D. L. (1988) *Proc. Natl. Acad. Sci. USA* **85,** 7036–7040.
22. Bisercic, M., Feutrier, J. Y. & Reeves, P. R. (1991) *J. Bacteriol.* **173,** 3894–3900.
23. Hall, B. G. & Sharp, P. M. (1992) *Mol. Biol. Evol.* **9,** 654–665.
24. Guttman, D. S. & Dykhuizen, D. E. (1994) *Science* **266,** 1380–1383.
25. Theisen, M., Borre, M., Mathiesen, M. J., Mikkelsen, B., Lebech, A.-M. & Hansen, K. (1995) *J. Bacteriol.* **177,** 3036–3044.
26. Thampapillai, G., Lan, R. & Reeves, P. (1994) *Mol. Biol. Evol.* **11,** 813–828.
27. Fiel, E., Carpenter, G. & Spratt, B. G. (1995) *Proc. Natl. Acad. Sci. USA* **92,** 10535–10539.
28. Li, W.-H., Wu, C.-I. & Luo, C.-C. (1986) *Mol. Biol. Evol.* **2,** 150–174.
29. Ochman, H., Whittam, T. S., Caugant, D. A. & Selander, R. K. (1983) *J. Gen. Microbiol.* **129,** 2715–2726.
30. Takiff, H. E., Salazar, L., Guerrero, C., Philipp, W., Huang, W. M., Kreiswirth, B., Cole, S. T., Jacobs, W. R., Jr. & Telenti, A. (1994) *Antimicrob. Agents Chemother.* **38,** 773–780.
31. Rouse, D. A., DeVito, J. A., Li, Z., Byer, H. & Morris, S. L. (1996) *Mol. Microbiol.* **22,** 583–592.
32. Kimura, M. (1983) *The Neutral Theory of Molecular Evolution* (Cambridge Univ. Press, Cambridge).
33. David, H. L. & Newman, C. M. (1971) *Am. Rev. Respir. Dis.* **104,** 508–515.
34. Bifani, P., Plikaytis, B. B., Kapur, V., Stockbauer, K., Pan, X., Lutfey, S. L., Moghazeh, S. L., Eisner, W., Daniel, T. M., Kaplan, M. H., Crawford, J. T., Musser, J. M. & Kreiswirth, B. N. (1996) *J. Am. Med. Assoc.* **275,** 452–457.
35. Collins, D. M., Kawakami, R. P., De Lisle, G. W., Pascopella, L., Bloom, B. R. & Jacobs, W. R., Jr. (1995) *Proc. Natl. Acad. Sci. USA* **92,** 8036–8040.
36. Kawaoka, Y. & Webster, R. G. (1988) *Microb. Pathog.* **5,** 311–318.
37. Simonet, M., Riot, B., Fortineau, N. & Berche, P. (1996) *Infect. Immun.* **64,** 375–379.
38. Hammerschmidt, S., Hilse, R., van Putten, J. P. M., Gerardy-Schahn, R., Unkmeir, A. & Frosch, M. (1996) *EMBO J.* **15,** 192–198.
39. Blower, S. M., Small, P. C. & Hopewell, P. C. (1996) *Science* **273,** 497–500.
40. Philipp, W. J., Poulet, S., Eiglmeier, K., Pascopella, Balasubramanian, V., Heym, B., Bergh, S., Bloom, B. R., Jacobs, W. R., Jr. & Cole, S. T. (1996) *Proc. Natl. Acad. Sci. USA* **93,** 3132–3137.
41. Govoni, G., Vidal, S., Gauthier, S., Skamene, E., Malo, D. & Gros, P. (1996) *Infect. Immun.* **64,** 2923–2929.